DOCUMENT RESUME

ED 299 898 HE 021 867

TITLE	A Proposal for Program Assessment at Kean College of New Jersey. Final Report of the Presidential Task Force on Student Learning and Development.
INSTITUTION	Kean Coll. of New Jersey, Union.
PUB DATE	May 86
NOTE	37p.
PUB TYPE	Information Analyses (070) -- Viewpoints (120)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Educational Assessment; Evaluation Criteria; *Evaluation Methods; Higher Education; National Surveys; Policy Formation; *Program Evaluation; Public Colleges; *School Policy; *Student Evaluation
IDENTIFIERS	*College Outcomes Assessment; *Kean College of New Jersey

ABSTRACT

In response to a request from its president, an institutional task force at Kean College of New Jersey studied approaches to program assessment used at other colleges, identified the measurement approaches being used at Kean College, and made policy recommendations for implementing an assessment program at the college. A major finding was that no single approach to assessment commands universal acceptance, but each institution conceives of assessment in ways it considers best suited to its unique needs and circumstances. The task force's final recommendations, derived from subcommittee reports, are presented in two groups. The first, designated as assessment guidelines, proposes several substantive principles to guide the assessment of learning at the college. The second group, designated as assessment adminstration, identifies specific actions to be taken to initiate and administer the college's assessment program. Appended materials include a statement of principles of the task force and a list of task force participants. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

A Proposal for Program Assessment at Kean College of New Jersey

Final Report of the Presidential Task Force on Student Learning and Development

ED299898

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

NATHAN WEISS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Prepared for

Dr. Nathan Weiss, President
Kean College of New Jersey
Union, N.J. 07083

May 1986

HE 021 867

KEAN COLLEGE OF NEW JERSEY
Union, New Jersey 07083

TABLE OF CONTENTS

EXECUTIVE SUMMARY 3

CONCLUSIONS AND RECOMMENDATIONS 11

Assessment Guidelines 12

Assessment Administration 31

REFERENCES 33

APPENDIX A—STATEMENT OF PRINCIPLES 34

APPENDIX B—TASK FORCE MEMBERSHIP 35

EXECUTIVE SUMMARY

This report by the Presidential Task Force on Student Learning and Development responds to the charge from Dr. Nathan Weiss to: 1) study approaches to program assessment which are being used at other colleges, 2) identify those measurement approaches now being used at Kean College, 3) disseminate information about these approaches to the Kean College community, 4) make policy recommendations for implementing an assessment program at Kean College, and 5) ensure that appropriate decision-making channels are used to approve recommendations.

The Task Force approached its work through the following four subcommittees: 1) National Approaches to Assessing Major Academic Programs, 2) National Approaches to Assessing General Education Programs, 3) Current Assessment Practices at Kean College, and 4) Approaches to Assessment: Existing Instruments and Measurement Criteria. Subcommittee members conducted their research through literature reviews, conference and seminar attendance, visits to other campuses, telephone calls, and surveys distributed on campus and mailed to other institutions.

Although the subcommittees identified many colleges and universities which engage in program assessment to one degree or another, only three institutions were found to have comprehensive assessment procedures in operation. These are Alverno College in Wisconsin, Northeast Missouri State University, and the University of Tennessee at Knoxville. Assessment at Alverno College differs significantly from assessment at the other two schools. At Alverno, individual students are assessed throughout the college experience with regard to their progress toward mastery of eight specific competencies. At Northeast Missouri and Tennessee, academic programs are assessed by measuring students' achievement levels at the conclusion of their courses of study.

A major finding of the subcommittees was that no single approach to assessment commands universal acceptance. Rather, each institution has conceived of assessment in ways which it believes to be best suited to its unique needs and circumstances. The moral for Kean College seems clear. Our assessment program should be consistent with the existing characteristics and commitments of our College.

The process of self-examination required by our assessment effort should be motivated by one central goal: to improve student learning. The steps in the process are: a) to articulate the goals of each academic program, b) to evaluate our students' progress toward those goals, and c) to use the resulting information to continue, modify or reinforce aspects of each academic program to ensure students' achievement. Kean College enters the assessment process from a position of confidence in its academic and cocurricular offerings, and views this as an opportunity to utilize additional approaches to assess the knowledge, capacities, skills, and values developed by our students.

In taking the opportunity to create and implement our own assessment

process, Kean College is seeking to develop what Alexander W. Astin (1984) calls "A new approach to excellence." Astin writes:

> Traditional measures of excellence—the reputational and resource approaches—clearly do not offer much hope of enhancing educational excellence or educational equity in our schools and colleges. . . . As an alternative to traditional approaches to defining educational excellence, I would like to suggest an approach which emphasizes educational impact . . . (p. 26).

As a significant part of Kean's three-year strategic plan, *Excellence and Equity*, assessment of student learning and development in all of our academic programs will provide useful guidelines for the College's continued fulfillment of our commitment to our students. The fact that the process will be designed and implemented by those who know that commitment best will ensure its consistency with the characteristics of the College.

The final recommendations of the Task Force emerged from the reports of the subcommittees. These recommendations are listed below. They have been arranged in two groups. The first group, designated Assessment Guidelines, proposes several substantive principles to guide the assessment of student learning at Kean College. The second group, designated Assessment Administration, identifies certain specific actions which should be taken to initiate and administer the College's assessment process. While the second group of recommendations is largely self-explanatory, a further word or two about the first group may be helpful.

The recommended Assessment Guidelines, in conjunction with their several subparts, are the result of careful thought and extensive discussion within the Task Force. The principal Task Force decisions which were the fruit of these deliberations were as follows: *First*, the Task Force decided to recommend that Kean College's assessment process should rely primarily on criterion-referenced, rather than norm-referenced, instruments. That principle is set forth in Recommendation Three below. *Second*, the Task Force decided to recommend that, regardless of the specific assessment strategy adopted, all programs at Kean College should clearly specify learning objectives for each course and should utilize valid and reliable assessment instruments in each course. That principle is set forth in Recommendation Four below. *Third*, the Task Force decided to recommend that there be a choice among five different assessment strategies or methods, rather than a uniform, College-wide approach. The five methods, which reflect the Task Force's awareness of the diversity of academic programs at the College, include three distinct strategies for conducting assessment at or near the end of students' work in programs and two distinct strategies for conducting assessment at earlier points in time. The five methods are set forth in Recommendation Five below. It should be noted that any one of these methods may be supplemented by another to complete a comprehensive program assessment.

A further point that deserves emphasis concerns the Task Force's recognition, expressed in Recommendation One below, that assessment involves im-

portant exploratory elements. Much is expected to be learned about the operation of the curriculum and its effects upon students as our assessment program evolves and undergoes development. Indeed, there is every reason to believe that many important assessment issues will only be discovered, and can only be addressed, after the process gets underway.

In view of the somewhat technical character of the considerations which went into the Task Force's decisions regarding assessment guidelines, a supporting narrative has been provided for these proposals. This narrative may be found in the Conclusions and Recommendations. (The page numbers in parentheses indicate the location of the supporting narrative for each recommendation.)

Assessment Guidelines

1. Assessment of student learning at Kean College should be approached as an exploration of our curriculum and the learning process associated with it (p. 12).

2. Our assessment program should be consistent with the mission of the College; it should reflect the diversity of our programs and the specific characteristics of our students' career objectives (p. 13).

3. Kean College should use criterion-referenced instruments in its assessment program. Primary responsibility for determining the appropriateness of such instruments should reside with the faculty in each program. The College may cooperate with testing organizations that might assist Kean College faculty in developing such instruments (p. 14).

   Two exceptions to this guideline are possible. (These exceptions may only be used to supplement the broader, comprehensive assessment program chosen by the faculty in accordance with Recommendation 5 below.):

   3a. While Kean College should not use norm-referenced instruments in its assessment program, this restriction does not preclude the use, as a supplement to the broader assessment process, of information derived from student performance on licensing examinations (for example, through an analysis of content subareas), provided that the passing rate on such examinations is not identified as an assessment of the program (p. 16).

   3b. While Kean College should not require the taking of post-baccalaureate admissions tests, such as the Graduate Record Examination, as part of its assessment program, insofar as students choose to take such tests for purposes of gaining admission to graduate or professional school information derived from analyses of content subareas of such tests may be used as a supplement to the broader assessment program (p. 18).

4. Assessment strategies adopted by the faculty should be based upon clear specification of learning objectives for each course and careful selection or design of valid and reliable assessment instruments (p. 20).

   4a. Assessment within the major should stress the development of skills related to the major, the development of human and professional values, and the knowledge of subject matter. The relationship among these three objectives should be established by each program's faculty (p. 22).

   4b. Each course in the curriculum should have well-defined objectives, including a specification of the student behaviors or performances sufficient to satisfy the objectives (p. 23).

4c. Multiple-choice examinations, if used, should be combined with additional assessment tools, such as written, oral or behavioral exercises (p. 25).

4d. Faculty members should be trained in the techniques of validating assessment instruments (p. 26).

5. Different methods of assessment may be appropriate for different programs. The Task Force recommends that the method of assessment for each program be selected from among the following five alternatives (p. 28).

5a. Programs in which students are required to take examinations for licensing or entry into a profession may use the test performances through content subarea analysis of the scores. Passing rates, however, may not be used as measures of program effectiveness. In addition, as proposed in Recommendation 3a, above, these programs will be expected to use faculty-developed assessment instruments to expand and enrich the assessment process for the particular discipline.

5b. Program assessment may utilize faculty-developed, criterion-referenced, end-of-program measurement instruments. Such instruments should provide for comprehensive and integrative indicators of the array of student competencies developed during participation in the program. Several approaches to end-of-program assessment may be appropriate including tests, performances, and portfolio evaluations. Student performance on the assessment measures may not, however, serve as a gate to prevent graduation.

5c. Program assessment may be conducted in the context of a synthesizing, credit-bearing senior seminar. This approach could enhance the curriculum and encourage learning while serving as a vehicle for an integrative, faculty-developed program assessment. Student performance on the assessment measures may not, by itself, prevent a student from receiving a passing grade in the seminar.

5d. Program assessment may be conducted through a series of "coherent clusters" within the discipline (e.g., a finance cluster in management science, a calculus cluster in mathematics). Measurement of achievement would take place by using criterion-referenced instruments in the cluster's culminating course. The cluster assessments would, as a series, provide for a comprehensive view of the program. Student performance on the assessment measures may not, by itself, prevent a student from passing the cluster's culminating course.

5e. Program assessment may be conducted within a series of key courses in a discipline. Measurement of achievement would take

place by using criterion-referenced instruments developed from a programmatic perspective to assess for sequenced levels of competencies. Student performance on the assessment measures may not, by itself, prevent a student from passing the course(s).

Assessment Administration

6. The "Statement of Principles" prepared by the Task Force and attached as Appendix A shall guide Kean College's assessment program.

7. Responsibility for developing an assessment program for general education should be assigned to a group including the Director of General Education, and other faculty members selected from the General Education Committee, from core course coordinators, and from faculty teaching general education courses. The assessment methods selected should be consistent with the recommendations, principles, and procedures in this report.

8. The policy for program review which was developed by the Faculty Senate and is being used at the College should be amended by adding the following:

    a) Faculty in each program should identify the educational objectives of the program.

    b) As a part of each program's review, measures of student learning, student satisfaction, and alumni responses should be used to assess the program's effectiveness. Other indicators of achievement which the program's faculty select may also be employed to assess the program.

9. A Kean College Program Assessment Council shall be established. The Council will be chaired by the President of the College, and the membership will include equal numbers of faculty members and administrators plus a representative from each of the student governing groups. The Senate Chairperson, the President of the local faculty bargaining agent, and the College President shall consult prior to designating Council members in order to help ensure appropriate representation. The Council will be charged to:

    a) Review the progress in implementing assessment programs.

    b) Review assessment procedures for consistency with guidelines proposed by the Faculty Senate and adopted by the Board of Trustees.

    c) Ensure that assessment information not be misused. Items 4, 5 and 6 in the attached Statement of Principles identify potential misuses of assessment data.

    d) Ensure that students are not tested excessively.

    e) Ensure that acceptable standards of sampling are employed when all students within a program are not involved in the assessment process.

f) Monitor the effects of assessment on all sectors of the student body, including foreign students, bilingual students, minority students, and part-time as well as other non-traditional students.

10. A program must be provided to develop faculty expertise in implementing effective assessment programs. The training must be directed at increasing effectiveness in: 1) setting appropriate goals and objectives, 2) preparing valid assessment instruments, 3) analyzing and interpreting assessment data, and 4) generating other knowledge and skills deemed helpful as the program progresses.

11. Responsibility must be assigned for coordinating the College's assessment program. Assessment coordinator(s) will be designated and they will be responsible for facilitating training opportunities, assisting faculty in developing assessment approaches, and integrating the assessment process with ongoing program review.

12. All aspects of assessing student learning and development should be coordinated with assessment of student life to provide for a comprehensive assessment approach by the College. The coordination process will be the responsibility of the coordinator(s) identified in Recommendation 11, above, and the Program Assessment Council.

CONCLUSIONS AND RECOMMENDATIONS

The Task Force has reviewed extensive literature on assessment of student learning and development and has examined methods of assessment at different institutions. No single definition of assessment is universally accepted, but the Task Force believes that assessment essentially involves self-definition and self-examination as steps toward self-improvement.

The goals of assessment are to validate the effectiveness of programs and, where appropriate, to make changes in the teaching-learning process. Peter Ewell (1984) identifies these goals as steps in the route to institutional excellence (p. 9). The Task Force recognizes that these are the only legitimate objectives for program assessment.

The Task Force further believes that no program should rely on a single measure to gain information about its effectiveness. A comprehensive approach to program review should include surveys of student and alumni satisfaction and achievement as well as indicators of learning outcomes.

The Task Force identified only three colleges and universities (Alverno College in Milwaukee, Wisconsin, Northeast Missouri State University, and the University of Tennessee at Knoxville) with comprehensive assessment programs in place. Kean College is in a position to be a leader in the area of student assessment. The efforts here are being observed closely by institutions and agencies across the nation.

There are several reasons why Kean College is well-suited to develop an effective assessment program: 1) The decision to begin such a program was made at the institution; it was not imposed by an outside agency. 2) Funding to support the project is available through the New Jersey Governor's Challenge Grant. 3) The College has engaged in extensive research and is in a position to benefit from the experiences of other institutions.

The Task Force recognizes that the assessment program implemented at Kean College must be designed to meet the goals of this institution and its diverse programs. Efforts by outside agencies to establish prescribed requirements or to standardize the process at the College would jeopardize the integrity of its individual programs and threaten the academic freedom of the faculty.

As Russell Edgerton, President of the American Association of Higher Education (1986) notes, "The real question is not whether assessment will arrive, but what shape it will take. Will we have simply a new overlay of standardized tests, or the kind of assessment that really improves learning (p. 11)?" The Task Force believes that the following recommendations will lead Kean College to a program that really improves learning.

RECOMMENDATIONS

Assessment Guidelines

1. ASSESSMENT OF STUDENT LEARNING AT KEAN COLLEGE SHOULD BE APPROACHED AS AN EXPLORATION OF OUR CURRICULUM AND THE LEARNING PROCESS ASSOCIATED WITH IT.

Assessment of Student Learning

This expression is preferable to the widely used catch-phrase "value-added assessment." That phrase derives from the discipline of economics where it denotes the monetary or cash value which accrues to raw materials as they pass through a production process. As Peter Ewell (1983) has pointed out: "The notion of *value-added* derived from this [economic] model . . . [e]ssentially . . . involves acceptance of the production process analogy (p. 7)." The effect of that analogy is that we are predisposed to value the results of the process rather than the process itself. Although that bias undoubtedly makes sense in the case of many economic processes, it has serious shortcomings when applied to education, especially at the college level where what is learned is not more important than the process of learning it.

For example, there may be no measurable difference in terms of *what* was learned between the student who knows certain things because her teachers told her so and the student who knows them because her teachers structured a situation in which she was led to find them out for herself. Nevertheless, to build a college-level curriculum exclusively around the former approach would be of dubious value. Even if it might result in greater gains in the short run in terms of what is learned (possibly because it is less time-consuming than the other approach), in the long run its graduates would be likely to suffer rapid obsolescence. Higher education at its best aims to create autonomous learners—self-directed, inquiring minds—and it is doubtful that such an objective can be achieved except by attending to process as well as to results (or, alternatively, means as well as ends), regardless of whether we define "results" as knowledge, skills, or values.

From this point of view, what is of fundamental importance in higher education is not just *how much* students have learned, but equally *how* students have learned. (Other similarly important questions include: How well have students learned? and How worthwhile to students is what they have been led to learn?) In sum, results are important and should be achieved, but to expect only that from college education is to expect too little. The expression "assessment of student learning" captures both the concern with process and the concern with results, and is preferable for that reason to the phrase "value-added assessment."

The term "assessment," when used with reference to college education, has no single, universally accepted definition. Advocates and practitioners of assessment in higher education are divided in their approaches to assessment by

their allegiance to distinct concepts of post-secondary learning. Two such divergent concepts appear to be of particular importance, namely, the result-oriented approach and the process-oriented approach. "Value-added assessment," as practiced by Northeast Missouri State University, typifies the result-oriented emphasis; "assessment," as practiced by Alverno College, exemplifies the emphasis upon process.

An Exploration of Our Curriculum and the Learning Process Associated With It

In terms of traditional criteria, the Kean College curriculum is achieving desirable goals. However, we presently have no agreed-upon way of understanding and describing its operation in detail. Assessment of student learning at the College should be thought of to begin with as a process of collecting information and developing a vocabulary, concepts, even a theory by means of which we can identify issues and begin to discuss them in a fashion which suits our concerns and our values. We need to know exactly what is happening before we can decide whether to reinforce it or to modify it.

For instance, we lack systematic information about the ways in which our students select or are recruited to particular majors, nor do we have substantial knowledge about the compatibility between our students and their majors. Yet compatibility between students and their programs of study would seem, on the face of it, as crucial to the success of the educational enterprise as compatibility of spouses is to the success of marriage.

Assessment in the sense of evaluation presupposes accurate identification and analysis of many such issues. The example of those who have preceded us in this undertaking indicates that it is a long-term process best carried out with utmost care and thoughtfulness.

2. OUR ASSESSMENT PROGRAM SHOULD BE CONSISTENT WITH THE MISSION OF THE COLLEGE; IT SHOULD REFLECT THE DIVERSITY OF OUR PROGRAMS AND THE SPECIFIC CHARACTERISTICS OF OUR STUDENTS' CAREER OBJECTIVES.

The Mission of the College

As both our Mission Statement and our strategic three-year plan emphasize, Kean College is committed to a policy of public access to a variety of baccalaureate degree programs. Within the limits of our resources, we are dedicated to serving the educational needs of all who can make good use of our services, regardless of age, sex, race, ethnic heritage, or previous history of economic, physical, or educational disadvantage. Moreover, Kean College is committed not merely to being accessible to previously disadvantaged or non-traditional groups of students, we are also committed to supporting and assisting "each student's intellectual and personal development" between admission and graduation

(*Excellence and Equity*, 1985, p. 6). That commitment has been carried out with a high degree of effectiveness of which all at the College may be proud. The specific features of our assessment program should be carefully chosen so that they enhance our continuing ability to deliver the benefits of higher education to all those who need and want them.

The Diversity of Our Programs and the Specific Characteristics of Our Students' Career Objectives

Registered nurses who come to the College for the purpose of acquiring the B.S.N. have quite different needs than do underprepared first-time freshmen just out of high school. Neither of those types of student is like the older citizen sampling course offerings for purposes of cultural enrichment but not yet decided whether to be committed to a degree program. Further, none of those three is in the same category as the part-time student enrolling for a few credits in the evening in order to improve chances for a promotion back on the job, let alone like the full-time overachiever bent from the first day on graduating with highest honors as a prelude to graduate or professional school. Yet the College has all of these types of students, and many more.

As our strategic three-year plan notes, our student population presents "a multitude of educational needs (p. 11)." Moreover, not only do we have this great diversity of students, we have also developed or are developing programs well-suited for each of them. Given this variety of students and programs, the task of designing an assessment process requires much more care and consideration than might otherwise be the case.

Our assessment program must reflect the diversity of our College. At the same time, we must not lose sight of the fact that we are a *college*, not merely a heterogeneous assortment of unrelated programs and individuals. With few exceptions every student is exposed to a wide range of approaches to undergraduate education, and even declared majors may decide to change their minds and major elsewhere. To a very great extent, the education of each student is a collective enterprise of the whole faculty and staff; the methods chosen by those in one program to assess "their" students' learning affect the interests of other programs, for in reality they are all our students.

The considerations set forth in the preceding paragraphs suggest several basic principles which should govern the choice of assessment instruments and techniques at Kean College.

3. KEAN COLLEGE SHOULD USE CRITERION-REFERENCED INSTRUMENTS IN ITS ASSESSMENT PROGRAM. PRIMARY RESPONSIBILITY FOR DETERMINING THE APPROPRIATENESS OF SUCH INSTRUMENTS SHOULD RESIDE WITH THE FACULTY IN EACH PROGRAM. THE COLLEGE MAY COOPERATE WITH TESTING ORGANIZATIONS THAT MIGHT ASSIST KEAN COLLEGE FACULTY IN DEVELOPING SUCH INSTRUMENTS.

Kean College Should Use Criterion-Referenced Instruments

The literature on assessment programs is filled with discussions of instrumentation. If the tester wants to learn how a group of students performs in comparison with students across the country, then "norm-referenced" instruments are appropriate. If the tester wants to determine whether students are learning what the instructor is trying to teach them, then the appropriate type of instrument to use is called "criterion-referenced." In more precise terms, the following definitions may be found useful:

> Criterion-Referenced Measurement: Measurement designed to assess an individual's status with respect to a particular criterion or standard of performance, irrespective of the relationship of his performance to that of others.
>
> Norm-Referenced Measurement: Measurement designed to assess an individual's standing with respect to other individuals on the same measuring device (Popham, 1973, pp. 98-99).

An "instrument," of course, is a measuring device. In education, any kind of question or instruction calling for some performance on the part of the student, where such performance is viewed as evidence of the student's possession of some desired knowledge, skill, or value is an instrument.

Norm-referenced instruments preclude, by design, equivalent performances by all those to whom they are administered, whereas criterion-referenced instruments do not. On a criterion-referenced test everyone might achieve desired levels of performance. A norm-referenced test, however, is designed to make comparisons among individual performances.

Familiar examples of criterion-referenced tests are typing tests and driving tests. So too, teacher-made tests are typically criterion-referenced. A teacher who shows a student how to dissect a laboratory animal, and who then presents the student with a fresh specimen accompanied by the instruction to dissect it, is using a criterion-referenced instrument to evaluate the student's possession of that skill. The criterion or standard of satisfactory performance is, of course, defined by the teacher.

Since Kean College is interested in determining whether its students benefit from the educational activities and resources made available to them, and is attempting to define what its students have learned, criterion-referenced instruments are the most appropriate type of instruments to use in our assessment program.

Determination of Instrument Appropriateness Should Reside With the Faculty in Each Program

This follows directly from the considerations in the preceding paragraphs. Assessment should be based on what students are being taught by their instructors. It is the faculty in each program who know best what the criteria or

standards of performance are relative to what they are attempting to teach. This is not a license to impose purely personal and subjective conceptions of such criteria since every responsible instructor is a representative of a discipline having its own internal structure and relatively objective standards of value.

3. (CONTINUED) TWO EXCEPTIONS TO THIS GUIDELINE ARE POSSIBLE. (THESE EXCEPTIONS MAY ONLY BE USED TO SUPPLEMENT THE BROADER, COMPREHENSIVE ASSESSMENT PROGRAM CHOSEN BY THE FACULTY IN ACCORDANCE WITH RECOMMENDATION 5 BELOW.):

   3a. WHILE KEAN COLLEGE SHOULD NOT USE NORM-REFERENCED INSTRUMENTS IN ITS ASSESSMENT PROGRAM, THIS RESTRICTION DOES NOT PRECLUDE THE USE, AS A SUPPLEMENT TO THE BROADER ASSESSMENT PROCESS, OF INFORMATION DERIVED FROM STUDENT PERFORMANCE ON LICENSING EXAMINATIONS (FOR EXAMPLE, THROUGH AN ANALYSIS OF CONTENT SUB-AREAS), PROVIDED THAT THE PASSING RATE ON SUCH EXAMINATIONS IS NOT IDENTIFIED AS AN ASSESSMENT OF THE PROGRAM.

Kean College Should Not Use Norm-Referenced Instruments

Norm-referenced testing involves "establishing norms for performance so that test scores come out as percentiles," where the higher the percentile, the better the score (Lazarus, 1977, pp. 186-187). Thus, a norm-referenced measurement of height or weight, for example, will not reveal, insofar as the results are stated as percentiles, how tall a person is or how much they weigh. Rather, it reveals what percentage of the measured population is shorter or taller, lighter or heavier, than the person in question. What this means is that test instruments of this type—for example, the College Outcomes Measures Program (ACT/COMP) and the Graduate Record Examination (GRE)—are designed so that they will reliably and predictably distribute student performances along a "normal" or bell-shaped curve.

The fundamental purpose of education should not be to teach students to test at higher levels than other students; it should be to impart desired levels of knowledge and to encourage certain abilities and values. A norm-referenced test, therefore, does not measure what Kean College's programs are designed to impart. A criterion-referenced instrument, because it reveals what a student has learned, does provide such information.

Moreover, whatever the merits of the preference for bell-shaped curves on the part of the makers and users of such tests (and there is an extensive literature debating the issue), the reality is that instruments of this type are subject to well-known abuses which make them unsuitable for use in an assessment program at Kean College. It is well established that performance on such tests is strongly

correlated with factors such as sex, socio-economic status, and race. While no informed consensus exists regarding the precise explanation of those correlations (the debate continues as vigorously as ever between those who blame the schools, those who blame the students, and those who blame the tests), that lack of consensus has not prevented certain sectors of public opinion from jumping to unwarranted negative conclusions about the relative educability of some groups of the kind which Kean College is committed to serving, for instance, those from low-income families. The consequences of using such instruments are predictable and would be counter-productive.

This Restriction Does Not Preclude The Use, As A Supplement To The Broader Assessment Process, Of Information Derived From Student Performance On Licensing Examinations (For Example Through An Analysis Of Content Subareas)

Many students at Kean College are required to pass licensing examinations for purposes of admission to certain occupations and professions. For example, all students intending to teach in the elementary and secondary schools must now pass the National Teacher Examination (NTE), a norm-referenced instrument. Recently, representatives of the company which manufactures that test, Educational Testing Service (ETS) of Princeton, New Jersey, met with Kean College faculty and administrators in order to familiarize them with the characteristics of the NTE and to identify, where possible, areas in which the College curriculum, or particular courses and course requirements, might be altered so as to help students who must take that test. The restriction on the use of norm-referenced instruments is not intended to preclude either the acquisition or the use of such obviously beneficial information.

Provided That The Passing Rate on Such Examinations is Not Identified As An Assessment of the Program

Programs may be structured with no intention of gearing curriculum to student passage of licensing examinations. The faculty may regard the development of certain skills, the encouragement of certain ways of thinking, the adoption of certain attitudes and values, and the accumulation of certain kinds of knowledge as more significant for the successful and appropriate practice of a certain profession than the mastery of the particular material that is tested on a licensing examination. The faculty may, moreover, expect students to prepare for licensing examinations through means that are external to the program itself (for example, commercial cram courses). The faculty should not necessarily be encouraged to harness programs to licensing examinations.

Additionally, pass rates on licensing examinations reflect two variables over which the College has little or no influence. One is the use of such examinations by licensing authorities to restrict access to certain occupations and professions. Pass rates on some licensing examinations have been observed to decline as the

supply of potential practitioners exceeded the number deemed desirable by various professional groups. The other is that individual student performance is likely to be strongly correlated with the intensity and self-discipline of a student's preparation to take the examination. Preparation for such an examination may involve a lengthy period of concentrated study, a great deal of memorization, and regular practice test-taking in order to increase the speed and accuracy with which time-pressured multiple-choice questions can be answered. Whether a student passes such an examination often depends on his or her determination to go through the tedious process of honing skills and marshalling knowledge into hair-trigger readiness. The College can provide skills and knowledge. But the College cannot be responsible for the determination needed to use them in the way required by these kinds of examinations. Pass rates on licensing examinations are, therefore, unreliable indices of the effectiveness of academic programs.

3b. WHILE KEAN COLLEGE SHOULD NOT REQUIRE THE TAKING OF POST-BACCALAUREATE ADMISSIONS TESTS, SUCH AS THE GRADUATE RECORD EXAMINATION, AS PART OF ITS ASSESSMENT PROGRAM, INSOFAR AS STUDENTS CHOOSE TO TAKE SUCH TESTS FOR PURPOSES OF GAINING ADMISSION TO GRADUATE OR PROFESSIONAL SCHOOL INFORMATION DERIVED FROM ANALYSES OF CONTENT SUBAREAS OF SUCH TESTS MAY BE USED AS A SUPPLEMENT TO THE BROADER ASSESSMENT PROGRAM.

Kean College Should Not Require The Taking Of Post-Baccalaureate Admissions Tests

In recent testimony before the National Governor's Association Task Force on College Quality in Washington, D.C., George H. Hanford (1986), President of the College Board, made the following comment about such tests:

> Such instruments as the Graduate Record Examination, the Graduate Management Admissions Test, the Law School Admissions Test, the National Teachers Examination, the Medical College Admissions Test, licensure tests for nursing or engineering, and so on . . . are sometimes *misused* to evaluate the quality of undergraduate programs just as SAT scores are sometimes *misused* to evaluate the quality of secondary schools. To the best of my knowledge there are no examinations specifically designed to assess what has happened to the basic academic competencies of college graduates during their undergraduate years . . . their abilities to communicate, deal with numbers, and reason (emphases added) (pp. 5-6).

Mr. Hanford's point is clear. The GRE and other similar instruments were designed to aid graduate and professional schools in making admissions decisions. To use them for any other purpose is to misuse them.

More specifically, the diversity of programs and career objectives typical of

Kean College students makes such instruments unsuitable for use in our assessment program. Most students at Kean College have not charted a career path leading through the particular levels and types of scholarly achievement prerequisite to graduate or professional training. It would be unsound, therefore, to assess our effectiveness in preparing students for the careers they have chosen by relying upon an instrument designed to predict success in post-baccalaureate training for careers they have not chosen. Ability to perform successfully on a graduate or professional school admission test is simply not the common measure of effective undergraduate education.

Moreover, there is a serious question of educational and professional ethics here which does not seem to have been widely recognized. Compelling students to take the GRE as a program assessment requirement may in some instances diminish their chances for later admission to graduate or professional school. Students who take the GRE exclusively to satisfy a program assessment requirement, rather than for purposes of post-baccalaurate admissions, are not exempt from the "cummulative score reporting" policy of the GRE Board. The GRE scores earned by such students at the time of graduation, along with any subsequent GRE scores, will automatically be reported by the Board to any graduate or professional schools which require GRE scores and to which the students apply during the three years following their college graduation. But if at the time those students took the GRE as a program assessment requirement they had no intention of applying to graduate or professional school, then they took the test under less-than-optimal conditions because the usual motivation to achieve a high score (namely, the desire to gain admission to graduate or professional school) was absent. If the graduate or professional schools to which such students subsequently apply (after retaking the test and improving their scores) have a policy of averaging an applicant's GRE scores, as many do and as the GRE Board recommends, then the students' chances of admission may be jeopardized by the earlier, poorer performance.

In short, to require of students not then intending to seek admission to graduate or professional school that they take such post-baccalaureate admissions tests as a program assessment requirement would be to place at risk their chances of such admission should they later decide to seek it—and it would be to do so without providing the students with any compensating benefit. Clearly, such a practice would be an unethical use of the College's power to set graduation requirements.

Except Insofar as Students Choose to Take Such Tests for Purposes of Gaining Admission to Graduate or Professional School

Where post-baccalaureate admissions tests have not been imposed upon students by the College as a program assessment requirement, and where students take such tests of their own free will because they want to go to graduate or professional school, then there is no reason why faculty who are interested in preparing students for such post-graduate education should not look to the test to glean whatever useful information they can from it (perhaps by content

analyses) for purposes of helping students in whatever way possible. As in the case of licensing exams, however, and for analogous reasons, student success rates should not be viewed as in any way constituting a program assessment.

The principal objective of programs from which students enter licensed professions or go on to graduate or professional schools is to prepare those students to perform well in those professions or in those schools, not to get them past the licensing or admissions tests (a job which commercial cram courses may be better equipped to perform). To become overly preoccupied with admission examinations, and to allow such examinations too much influence over curriculum decisions would be, in effect, to misconceive the purpose of undergraduate education, especially in view of the fact that there may often be little connection between the content of such examinations and the content of graduate and professional school curricula.

4. ASSESSMENT STRATEGIES ADOPTED BY THE FACULTY WITHIN THE PROGRAMS AT KEAN COLLEGE SHOULD BE BASED UPON CLEAR SPECIFICATION OF LEARNING OBJECTIVES FOR EACH COURSE AND CAREFUL SELECTION OR DESIGN OF VALID AND RELIABLE ASSESSMENT INSTRUMENTS.

However profound their differences in other respects, institutions which have undertaken the project of establishing some type of assessment process are in perfect agreement on one point: such a project requires the setting of objectives and the identification of means or measures by which progress toward the objectives can be reliably evidenced. Thus, Northeast Missouri State University (1984):

> The institution must specify its expectations for student learning, decide when its milestones are to be reached, and select appropriate measures (p. 25).

Thus, the University of Tennessee at Knoxville (Banta and Fisher, 1984):

> Comprehensive program evaluation . . . . . [proceeds] from the setting of objectives, through the allocation and utilization of resources for implementation, to the measurement of outcomes related to the objectives (p. 30).

Thus, Alverno College (Alverno College Faculty, 1985):

> The principle [behind assessment] is simple: to begin by being as clear as possible about the expected outcomes. . . . [W]e have found that we cannot begin to create an effective assessment situation until we are clear about *which* specific abilities we are looking to see demonstrated and *how* they might appear (p. 17).
>
> After we had identified these [expected outcomes], the question became: 'How can we tell how far along a student is in developing these competencies?' It would be pointless, we realized, to have spent so much time and effort articulating our educational goals unless we were willing to make a similar investment in assessing their attainment (p. 6).

Differences appear, however, regarding method. For example, Northeast Missouri State University (1984) seems to propose a deductive or quasi-deductive "progression from [institutional] mission statement to specific learning objectives [for programs]," with no clear indication whether the deduction should proceed further to specify objectives for particular courses, course clusters or course sequences (pp. 17-18). On the other hand, while Alverno College (1985) also proceeded, in a sense, from a set of institutional objectives, namely, eight general abilities or competencies which it seeks to develop in all students, it also made use of a converse inductive or quasi-inductive procedure: "We began with individual faculty members and departments struggling to take abilities that were implicit as outcomes of their courses and turn them into something concrete and observable that students could be expected to do (p. 17)."

The present recommendation suggests that at Kean College the process of specifying assessment objectives should be developed inductively at the program level, starting with individual courses or course sequences. The rationale for this suggestion is well-expressed in the following piece of practical wisdom offered by Northeast Missouri State University (1984): "[A] prudent institution would want evidence that it is effective in developing some competency *before* it required all students to demonstrate it (p. 42)." (emphasis added). Or, to put the point a bit differently, we should not attempt to define our objectives until we determine what we can do well.

Of course, a purely inductive procedure would be inadequate. Course objectives must be defined with an eye on program objectives, for courses must fit together in some connected fashion suggested by the principles of the discipline. Yet program objectives formulated without regard for the realities of course offerings may become meaningless. As Immanuel Kant might have put it: course objectives without program objectives are blind, program objectives without course objectives are empty.

In the course of working their way through this dialectical dilemma, faculty in programs should attend to the following kinds of questions and issues which were raised by the Task Force Subcommittee on Assessment of Major Programs:

> When one speaks of assessment within major programs, which of the following does one mean?
>
> a. Assessing the effectiveness of the program?
>
> b. Assessing an individual student's knowledge of the subject matter?
>
> c. Assessing all students' knowledge of the subject matter?
>
> d. Assessing an individual student's skill competencies at a particular level?
>
> e. Assessing all students' skill competencies at a particular level?
>
> f. Helping the student to improve?
>
> g. Moving toward improvement of the curriculum?

Concerns regarding instrumentation have been addressed, in part, under Recommendation 3, above. The import of that recommendation is that Kean College will rely on locally developed assessment instruments, except insofar as content analyses of licensing and other such exams are useful, and except insofar as criterion-referenced instruments which have been developed commercially or by other colleges and universities become available.

The Sub-Recommendations below further address, in the case of 4a and 4b, the issue of assessment objectives, and, in the case of 4c and 4d, the issue of locally developed instrumentation.

4a. ASSESSMENT WITHIN THE MAJOR SHOULD STRESS THE DEVELOPMENT OF SKILLS RELATED TO THE MAJOR, THE DEVELOPMENT OF HUMAN AND PROFESSIONAL VALUES, AND THE KNOWLEDGE OF SUBJECT MATTER. THE RELATIONSHIP AMONG THESE THREE OBJECTIVES SHOULD BE CLEARLY ESTABLISHED BY EACH PROGRAM'S FACULTY.

Successful college education is not limited to the attainment of purely cognitive objectives. The baccalaureate degree does not simply signify that a graduate possesses some particular body of information. Rather, the degree is commonly understood to signify, first, that the graduate has achieved a certain mastery of the fundamentals of some discipline, and, second, that the graduate has benefitted from exposure to the broadening and deepening influence of the general education portion of the curriculum. The acquisition of information is a part of both of these components of learning at the college level more because it has served as a vehicle for developing the student's competence within a discipline and the student's general intellectual sophistication than because such acquisition is an end in itself.

These points are commonly expressed by saying that college education is at least as much concerned with skills and values as it is with knowledge of subject matter in the narrow sense. To give but one example, the physics graduate is not expected only to know what the results of past experiments have been, but also to have a well-developed skill in the technique of hypothesis formation, experimental design, and the interpretation of results, and, in addition, to have a profound commitment, among other things, to the values of honesty, precision, and accuracy in the reporting or publication of experimental data. Information, skills, and values form a seamless web in the acquisition of genuine competence in each discipline.

Moreover, the same physics graduate should have acquired, as part of the overall college experience, a well-articulated understanding of the uniqueness of physical science and its role in the history of human civilization. The physicist who does not understand why physics is not all there is to thought and to life is that much less of a physicist for failing to grasp the specific strengths and limits of the contributions of physics to the good of humanity.

In view of these considerations, academic departments should take particular care, as they review and define their objectives, to explicate and clarify the

interrelationships of the subject matter that is being taught, the skills which every student in that discipline should, as a specialist, develop during the course of study, and the human and professional values inherent in the discipline. In some disciplines, certain professional values are consciously formulated and taught, for example, in a professional ethics component of the curriculum. In others, the values implicit in the discipline or the curriculum may be less well or less often consciously articulated. Where that is so, it is suggested that special attention should be devoted to educating students in this critical aspect of college learning.

The knowledge and skills which students acquire in college are the elements of a precious heritage, and society is entitled to expect that the time and resources devoted to the education of the individual will translate into constructive contributions to the community. In that sense, to fail to communicate the values inherent in the enterprise of higher education is fundamentally to fail to educate.

4b. EACH COURSE IN THE CURRICULUM SHOULD HAVE WELL-DEFINED OBJECTIVES, INCLUDING A SPECIFICATION OF THE STUDENT BEHAVIORS OR PERFORMANCES SUFFICIENT TO SATISFY THE OBJECTIVES.

To varying degrees, most courses at the College already exhibit these features. Program objectives are commonly thought of as specific competencies which the successful student will make his or her own, and course objectives are typically conceived of, loosely, as steps on an ascending ladder.

The most important principle in the formulation of course objectives is stated well by Popham (1973) in the following passage:

> Most educators agree that educational objectives are important. They will readily agree that the teacher does really not know what to do without instructional goals, that he does not really know what to plan for in his classes without them. Unfortunately, however, most instructional objectives are stated so vaguely and so ambiguously that they are of limited value. When objectives are phrased in very broad terms, there are many possible interpretations that could be made about what they mean and it becomes almost impossible to figure out what the instructor really has in mind. For example, when a teacher says he wants his students to 'become familiar with' a certain concept, we do not know how they are supposed to manifest this familiarity unless the instructor goes on to state just what the students should be able to do in order to evidence the desired familiarity.
>
> Meaningful instructional objectives must be stated in terms of student behavior, and they must specify the type of behavior a student will engage in (or be able to engage in) when he has satisfactorily achieved the objective. For instance, the teacher might assert first in a rather general way that the student will 'understand' a particular principle, and then he might clarify this goal by giving a more specific objective. This objective could indicate, for example, that understanding will be manifested when the student is able

to describe in writing the five major elements of the principle and give one original example of each.

In other words, the more explicit the instructor can be regarding the statement of instructional objectives, the better. The only kind of specificity that really helps in improving teacher behavior empirically is the specification of goals in terms of student behavior changes. 'Behavior' in this sense is defined rather broadly. It does not need to be behavior only on a test; it may also include behavior in an informal situation (for example, observed politeness) or even behavior in a situation specifically contrived by the instructor to reflect a *more basic kind* of commitment. For instance, the literature instructor is often concerned with the degree to which his students are really interested in the literature he has been treating. He may set up situations in which students are free to choose reading material, and he may then note the degree to which their choices coincide with preferences advocated in his class. The behavior used most commonly to measure behaviorally stated objectives is test behavior, although there are numerous other ways to assess objectives in the intellectual as well as attitudinal realms. These latter approaches take considerable ingenuity on the part of the teacher (p. 13).

Although, as this passage indicates, specificity is a necessary quality of useful objectives, it must not be so extreme that the objective becomes, as it is called, a "test item equivalent" (for example, "The student will be able to list the names of five ancient Greek philosophers"), for in that case hundreds of objectives might be needed for each course. Rather, course objectives should ordinarily possess a certain degree of content generality, i.e., they should "describe behavior that can generalize across a range of content (Popham, 1973, p. 15)." (For example, "The student will be able to identify and explain in writing any sequence of two or more classical Greek thinkers who developed a shared philosophical principle.")

Then, too, minimally acceptable standards of student performance necessary to satisfy the objectives should be specified in advance and made apparent to the students at the beginning of the course or at the beginning of a discrete segment of it.

Although behavioral objectives are sometimes connected with an emphasis on objective or multiple-choice testing, the present Recommendation contemplates no such connection in the assessment of student learning at Kean College. As suggested by Recommendation 4c, below, test behavior may, in some learning situtations, be an inferior and unreliable method of assessing student learning, especially as compared with the direct observation of a student's performance of the specific types of intellectual, creative, or other work valued by each discipline. Moreover, as indicated in Recommendation 1, above, college-level learning objectives should be understood to encompass not only desired outcomes or results but also—and sometimes more importantly—the activities and processes by which those results are effected. The developing body of writing on college-level teaching produced by the faculty at Alverno College provides

an illuminating example of the definition of behavioral objectives in terms of both educational process and educational outcomes within the context of a pedagogy committed to avoiding the pitfalls of objective testing.

4c. MULTIPLE-CHOICE EXAMINATIONS, IF USED, SHOULD BE COMBINED WITH ADDITIONAL ASSESSMENT TOOLS, SUCH AS WRITTEN, ORAL, OR BEHAVIORAL EXERCISES.

Teachers who consider using multiple-choice questions face a dilemma. On the one hand, this testing format is disdained by many educators on the grounds that it is too easy. If students are given the correct answer, albeit along with several incorrect answers, then those who merely "know it when they see it" (not to mention those who simply guess well) have a relatively good chance of being credited with knowledge of the matter in question. Yet "knowing it when you see it," while a step up from the accurate guess, is arguably the lowest and most passive level of knowing and not—as such things are estimated among college and university educators—worth very much.

On the other hand, the multiple-choice question has been criticized on the grounds that it is too difficult. If students are to be presented in such questions with a real test of what they know, then the correct answer must be well-hidden among deceptively similar, but false, alternatives (indeed, the false alternatives are technically referred to in the literature on testing as "distractors"). In other words, the challenge represented by such questions is a function of the subtlety of the distinctions between the alternative answers, and with subtlety inevitably comes ambiguity. While the point in the preceding paragraph suggests that multiple-choice questions may reward the student who knows too little, this one argues that such questions may penalize the student who knows too much. Such students may see things in the question and its subtleties which were not intended to be there, and so be misled into selecting erroneous answers.

The foregoing considerations spring from concern with the issue of fairness in testing (rewarding the less knowledgeable versus penalizing the more knowledgeable). Other familiar criticisms of the multiple-choice format address some equally important concerns. For example, the trouble with multiple-choice questions from the perspective of the teacher concerned with measuring his or her effectiveness in the classroom is that they do not usually provide much opportunity for identifying why students make the selections they do. Consequently, such questions neither permit confirmation of what is most highly valued in college-level learning (namely, ability to give reasons for what one believes to be true) nor allow for diagnosis of what might have led or what might lead a particular student astray (such as good reasons for a wrong answer or bad reasons for a right answer).

Again, from the perspective of leading students toward genuine competence in a discipline, the multiple-choice format has the shortcoming that it tends to suggest to the student that knowledge consists of little bits and pieces of information and that inquiry is a matter of sorting these disconnected fragments into piles which can be labeled "correct" or "incorrect." Not only does the multiple-

choice examination typically fail to communicate the coherent character of a discipline, but the relative weight and significance of a discipline's component elements, the guiding objectives or purposes which lie behind it, and the inherently tentative and provisional character of its findings at any given point in time are also lost to view. One should not underestimate the impressions students form about the nature of a discipline from the methods which are used to determine what they have learned.

Fairness toward students, the classroom teacher's need to know why, the imperative that each discipline be conveyed to students in non-simplistic ways—all argue against extensive reliance on multiple-choice examinations as a means to assess college-level learning. Needless to say, more or less the same considerations tell against other methods of objective testing, such as true/false questions, fill-ins, matching exercises and so forth. In all of these cases, the ease and certainty with which student learning is measured is purchased by short-changing the student, the teacher, and the discipline.

Competence in each discipline involves integrating a unique combination of knowledge, skills, and values into a serviceable capacity to perform the type of intellectual or creative work for which the discipline exists. There is no substitute for the actual performance of such work as a means of determining whether students have acquired the capacity to perform it. Exercises involving oral or written communication or other behavioral performances of the type valued by the discipline provide both teachers and students with the most reliable means of verifying mastery of the essential competencies. For example, presentation of a portfolio, completion of a research experiment, or the writing and defending of a scholarly senior thesis are among the methods by which students may demonstrate not only understanding of subject matter but also the ability to integrate that knowledge with the skills and values they have developed in their discipline.

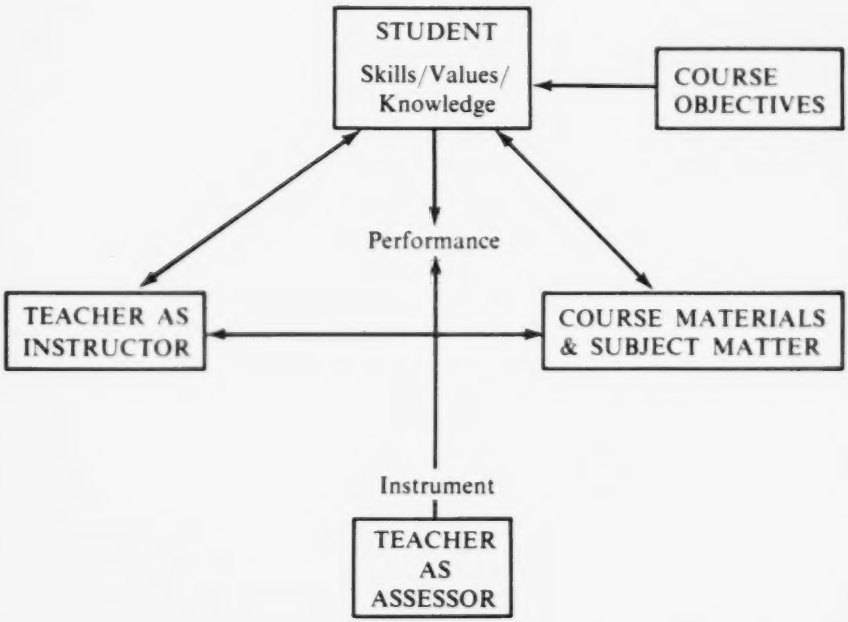
4d. FACULTY MEMBERS SHOULD BE TRAINED IN THE TECHNIQUES OF VALIDATING ASSESSMENT INSTRUMENTS.

Although every experienced faculty member has sought to improve the test instruments used in his or her courses, it does not follow that faculty members could not benefit from and would not welcome further experiences in this aspect of teaching. It must be remembered that most college faculty never received any formal training or guidance in teaching or testing techniques. As one experienced professor (Fraher, 1984) has put it: "Professors are trained only as scholars and then thrust in front of the classroom to play the role of teacher. To say that this transition in roles can be a learning experience is to indulge in understatement (p. 116)." Or, as Popham (1973) suggests: "Any person who sets out systematically to improve his teaching skill is embarking on a challenging but rewarding task. . . . Many *experienced* teachers come to realize that their approaches to instruction have been largely instinctual. Even seasoned veteran teachers can profit by approaching the instructional act with a degree of precision (p. 9)"

Test instruments are customarily evaluated in terms of their reliability and validity. Instrument reliability is a measure of consistency. An unreliable instrument is like an elastic yardstick which produces a different result each time it is used. Thus, assuming that practice makes no difference, if the same instrument is administered twice to the same group of students and produces a different result each time, then to that extent the instrument is thought to lack reliability. In practice, of course, classroom teachers can get only a rough sense of reliability by administering the same or similar instruments to different groups of students. Faculty sharing of instruments may enhance this process and may provide an important benefit to our students.

A valid instrument is one that measures what it is intended to measure. Jonathan Warren (1986) defines validity as a measure of the degree of empirical support which student responses to the instrument provide for the inference that the student has or does not have the desired knowledge, skill, or value.

Every teacher knows the principle: the questions you ask determine the answers you get. The theory which underlies the concept of instrument validity rests on the premise that a student's knowledge, skills, and values are unobservable entities. A teacher can determine whether they are present or not only by inference from something observable. What is observable are student responses (behaviors or performances) to a teacher's questions or instructions. The situation may be diagrammed in this way:



To be valid, an instrument should be designed to elicit behavior or performances which clearly manifest the knowledge, skills, or values which it is the objective of the course to teach, and which the course objectives have clearly specified.

Given the above definition of validity, the following propositions are true:

1) if the instrument is *valid,* then student performance/non-performance supports the inference that the student possesses/does not possess the skill, value or knowledge in question;

2) if the instrument is *not valid,* then student performance/non-performance does not warrant the inference that the student possesses/does not possess the desired skill, value or knowledge;

3) if the instrument is *valid,* then student non-performance supports the inference that "something is wrong" with the learning process, i.e., the three-cornered interaction among teacher, student, and course materials-subject matter; and

4) while students may be graded, i.e., ranked in some order of merit, using any instrument regardless of its validity, a *valid* instrument permits both grading and assessment—what is graded is the student, what is assessed is the learning process.

5. DIFFERENT METHODS OF ASSESSMENT MAY BE APPROPRIATE FOR DIFFERENT PROGRAMS. THE TASK FORCE RECOMMENDS THAT THE METHOD OF ASSESSMENT FOR EACH PROGRAM BE SELECTED FROM AMONG THE FOLLOWING FIVE ALTERNATIVES.

Presented below are five distinct methods for the assessment of student learning. These reflect the Task Force's awareness of the diversity of program characteristics at the College, and the recognition that different assessment strategies may be suitable for different programs. The Task Force believes that these five methods maximize the range of assessment options. Moreover, the Task Force believes that each strategy is equally capable of comprehensively assessing the student learning process. A few words about each of these points may be in order.

During the past two decades the mission of Kean College has changed dramatically. Once largely devoted to teacher education, the College now offers a comprehensive array of baccalaureate-level programs in the arts and sciences and in the professional disciplines. The differences among these programs in size, curricular characteristics, vocational orientation, and teaching staff—to mention only the most noticeable features—are substantial. Some programs have no more than a dozen or two dozen majors while others serve the needs of many hundreds. Some programs, especially many in the professional disciplines, have highly structured and tightly sequenced curricula with all students pursuing essentially the same required course of study, while in others students may pursue many

different paths through a very loosely structured set of course requirements. Some programs are closely tied to the standards of accrediting or other external agencies which regulate access to licensed occupations and professions, while others are responsive only to the diffuse and slowly changing currents in a scholarly or creative discipline with no specific connection to any particular career. Some programs use many adjunct instructors, while in most others virtually all teaching is conducted by the resident, full-time faculty.

The Task Force has not attempted to analyze in detail precisely how such differences may bear on decisions about assessment strategies. However, one example may be sufficient to indicate that different strategies might suit different programs, and therefore that the provision of alternatives is the wisest course. A program with a relatively small number of majors, with a high ratio of resident faculty to majors, and with sufficient flexibility in its existing course requirements might decide to focus its assessment program around a credit-bearing senior seminar as described in 5c below. On the other hand, a program with a very large number of majors, with a low ratio of resident faculty to majors, and with little flexibility in its course requirements due to external accrediting standards might find that particular option difficult to manage effectively. Consequently, depending on its other characteristics, the second program might find one of the other options better suited to its needs.

The Task Force believes that the alternative methods recommended are sufficient to accommodate the diverse needs of our programs. Thus, assessment may be focused either at the end of a student's work in a program (5b and 5c) or at earlier points in time while the student's work is still underway (5d and 5e). If assessment is focused at the end of the program, then it may be carried out either within the context of a course (5c) or outside the context of a course (5b). If assessment is focused at earlier points in time, then it may be tied in to course sequences or course clusters where they exist (5d) or to other key points in the course curriculum where they do not (5e). Alternative 5a is *sui generis*. It assumes that, in programs in which students take licensing examinations, content analyses of the results of those examinations will supplement an assessment strategy selected from the other four alternatives. Alternative 5a makes clear that licensing examinations alone are not sufficient to constitute a satisfactory assessment program.

Finally, each assessment strategy is intended to represent a possible way of comprehensively assessing the student learning process in the program or discipline. Comprehensiveness is understood in two senses. First, comprehensiveness refers to the assessment goals and objectives of the program or discipline, which should encompass the full range of knowledge, skills, and values which the faculty in that program or discipline seek to instill. Second, comprehensiveness refers to the extent to which assessment gathers information about the effect of a program on all students who are enrolled in it. It is recognized that some large programs may need to rely upon a sampling procedure, and that comprehensiveness in this second sense may be achieved in such cases insofar as the student sample selected is representative of all students enrolled in the program.

5a. PROGRAMS IN WHICH STUDENTS ARE REQUIRED TO TAKE EXAMINATIONS FOR LICENSING OR ENTRY INTO A PROFESSION MAY USE THE TEST PERFORMANCES THROUGH CONTENT SUBAREA ANALYSIS OF THE SCORES. PASSING RATES, HOWEVER, MAY NOT BE USED AS MEASURES OF PROGRAM EFFECTIVENESS. IN ADDITION, AS PROPOSED IN RECOMMENDATION 3A, ABOVE, THESE PROGRAMS WILL BE EXPECTED TO USE FACULTY-DEVELOPED ASSESSMENT INSTRUMENTS TO EXPAND AND ENRICH THE ASSESSMENT PROCESS FOR THE PARTICULAR DISCIPLINE.

5b. PROGRAM ASSESSMENT MAY UTILIZE FACULTY-DEVELOPED, CRITERION-REFERENCED, END-OF-PROGRAM MEASUREMENT INSTRUMENTS. SUCH INSTRUMENTS SHOULD PROVIDE FOR COMPREHENSIVE AND INTEGRATIVE INDICATORS OF THE ARRAY OF STUDENT COMPETENCIES DEVELOPED DURING PARTICIPATION IN THE PROGRAM. SEVERAL APPROACHES TO END-OF-PROGRAM ASSESSMENT MAY BE APPROPRIATE INCLUDING TESTS, PERFORMANCES, AND PORTFOLIO EVALUATIONS. STUDENT PERFORMANCE ON THE ASSESSMENT MEASURES MAY NOT, HOWEVER, SERVE AS A GATE TO PREVENT GRADUATION.

5c. PROGRAM ASSESSMENT MAY BE CONDUCTED IN THE CONTEXT OF A SYNTHESIZING, CREDIT-BEARING SENIOR SEMINAR. THIS APPROACH COULD ENHANCE THE CURRICULUM AND ENCOURAGE LEARNING WHILE SERVING AS A VEHICLE FOR AN INTEGRATIVE, FACULTY-DEVELOPED PROGRAM ASSESSMENT. STUDENT PERFORMANCE ON THE ASSESSMENT MEASURES MAY NOT, BY ITSELF, PREVENT A STUDENT FROM RECEIVING A PASSING GRADE IN THE SEMINAR.

5d. PROGRAM ASSESSMENT MAY BE CONDUCTED THROUGH A SERIES OF "COHERENT CLUSTERS" WITHIN THE DISCIPLINE (E.G., A FINANCE CLUSTER IN MANAGEMENT SCIENCE, A CALCULUS CLUSTER IN MATHEMATICS). MEASUREMENT OF ACHIEVEMENT WOULD TAKE PLACE BY USING CRITERION-REFERENCED INSTRUMENTS IN THE CLUSTER'S CULMINATING COURSE. THE CLUSTER ASSESSMENTS WOULD, AS A SERIES, PROVIDE FOR A COMPREHENSIVE VIEW OF THE PROGRAM. STUDENT PERFORMANCE ON THE ASSESSMENT MEASURES MAY NOT, BY ITSELF, PREVENT A STUDENT FROM PASSING THE CLUSTER'S CULMINATING COURSE.

5e. PROGRAM ASSESSMENT MAY BE CONDUCTED WITHIN A SERIES OF KEY COURSES IN A DISCIPLINE. MEASUREMENT OF ACHIEVEMENT WOULD TAKE PLACE BY USING CRITERION-REFERENCED INSTRUMENTS DEVELOPED FROM A PROGRAMMATIC PERSPECTIVE TO ASSESS FOR SEQUENCED LEVELS OF COMPETENCIES. STUDENT PERFORMANCE ON THE ASSESSMENT MEASURES MAY NOT, BY ITSELF, PREVENT A STUDENT FROM PASSING THE COURSE(S).

Assessment Administration

6. THE "STATEMENT OF PRINCIPLES" PREPARED BY THE TASK FORCE AND ATTACHED AS APPENDIX A TO THIS REPORT SHALL GUIDE KEAN COLLEGE'S ASSESSMENT PROGRAM.

7. RESPONSIBILITY FOR DEVELOPING AN ASSESSMENT PROGRAM FOR GENERAL EDUCATION SHOULD BE ASSIGNED TO A GROUP INCLUDING THE DIRECTOR OF GENERAL EDUCATION, AND OTHER FACULTY MEMBERS SELECTED FROM THE GENERAL EDUCATION COMMITTEE, FROM CORE COURSE COORDINATORS, AND FROM FACULTY TEACHING GENERAL EDUCATION COURSES. THE ASSESSMENT METHODS SELECTED SHOULD BE CONSISTENT WITH THE RECOMMENDATIONS, PRINCIPLES, AND PROCEDURES IN THIS REPORT.

8. THE POLICY FOR PROGRAM REVIEW WHICH WAS DEVELOPED BY THE FACULTY SENATE AND IS BEING USED AT THE COLLEGE SHOULD BE AMENDED BY ADDING THE FOLLOWING:

   a) FACULTY IN EACH PROGRAM SHOULD IDENTIFY THE EDUCATIONAL OBJECTIVES OF THE PROGRAM.

   b) AS A PART OF EACH PROGRAM'S REVIEW, MEASURES OF STUDENT LEARNING, STUDENT SATISFACTION, AND ALUMNI RESPONSES SHOULD BE USED TO ASSESS THE PROGRAM'S EFFECTIVENESS. OTHER INDICATORS OF ACHIEVEMENT WHICH THE PROGRAM'S FACULTY SELECT MAY ALSO BE EMPLOYED TO ASSESS THE PROGRAM.

9. A KEAN COLLEGE PROGRAM ASSESSMENT COUNCIL SHALL BE ESTABLISHED. THE COUNCIL WILL BE CHAIRED BY THE PRESIDENT OF THE COLLEGE AND THE MEMBERSHIP WILL INCLUDE EQUAL NUMBERS OF FACULTY MEMBERS AND ADMINISTRATORS PLUS A REPRESENTATIVE FROM EACH OF THE STUDENT GOVERNING GROUPS. THE SENATE CHAIRPERSON, THE PRESIDENT OF THE LOCAL FACULTY BARGAINING AGENT, AND THE COLLEGE PRESIDENT SHALL CONSULT PRIOR TO DESIGNATING COUNCIL MEMBERS IN ORDER TO HELP ENSURE AP-

PROPRIATE REPRESENTATION. THE COUNCIL WILL BE CHARGED TO:

a) REVIEW THE PROGRESS IN IMPLEMENTING ASSESSMENT PROGRAMS.

b) REVIEW ASSESSMENT PROCEDURES FOR CONSISTENCY WITH GUIDELINES PROPOSED BY THE FACULTY SENATE AND ADOPTED BY THE BOARD OF TRUSTEES.

c) ENSURE THAT ASSESSMENT INFORMATION NOT BE MISUSED. ITEMS 4, 5, AND 6 IN THE ATTACHED STATEMENT OF PRINCIPLES IDENTIFY POTENTIAL MISUSES OF ASSESSMENT DATA.

d) ENSURE THAT STUDENTS NOT BE TESTED EXCESSIVELY.

e) ENSURE THAT ACCEPTABLE STANDARDS OF SAMPLING ARE EMPLOYED WHEN ALL THE STUDENTS WITHIN A PROGRAM ARE NOT INVOLVED IN THE ASSESSMENT PROCESS.

f) MONITOR THE EFFECTS OF ASSESSMENT ON ALL SECTORS OF THE STUDENT BODY INCLUDING FOREIGN STUDENTS, BILINGUAL STUDENTS, MINORITY STUDENTS, AND PART-TIME AS WELL AS OTHER NON-TRADITIONAL STUDENTS.

10. A PROGRAM MUST BE PROVIDED TO DEVELOP FACULTY EXPERTISE IN IMPLEMENTING EFFECTIVE ASSESSMENT PROGRAMS. THE TRAINING MUST BE DIRECTED AT INCREASING EFFECTIVENESS IN: 1) SETTING APPROPRIATE GOALS AND OBJECTIVES, 2) PREPARING VALID ASSESSMENT INSTRUMENTS, 3) ANALYZING AND INTERPRETING ASSESSMENT DATA, AND 4) GENERATING OTHER KNOWLEDGE AND SKILLS DEEMED HELPFUL AS THE PROGRAM PROGRESSES.

11. RESPONSIBILITY MUST BE ASSIGNED FOR COORDINATING THE COLLEGE'S ASSESSMENT PROGRAM. ASSESSMENT COORDINATORS WILL BE DESIGNATED, AND THEY WILL BE RESPONSIBLE FOR FACILITATING TRAINING OPPORTUNITIES, ASSISTING FACULTY IN DEVELOPING ASSESSMENT APPROACHES, AND INTEGRATING THE ASSESSMENT PROCESS WITH REGULAR PROGRAM REVIEW.

12. ALL ASPECTS OF ASSESSING STUDENT LEARNING AND DEVELOPMENT SHOULD BE COORDINATED WITH ASSESSMENT OF STUDENT LIFE TO PROVIDE FOR A COMPREHENSIVE ASSESSMENT APPROACH BY THE COLLEGE. THE COORDINATION PROCESS WILL BE THE RESPONSIBILITY OF THE COORDINATOR(S) IDENTIFIED IN RECOMMENDATION 11, ABOVE, AND THE PROGRAM ASSESSMENT COUNCIL.

REFERENCES

Alverno College Faculty. *Assessment at Alverno College,* Rev. Edit. Milwaukee: Alverno College, 1985.

Astin, Alexander W. "Excellence and Equity: Achievable Goals for American Education," *Phi Kappa Phi Journal,* 64 (Spring 1984) 24-29.

Banta, Trudy W. and Fisher, Homer S. "Performance Funding: Tennessee's Experiment," in J. Folger (ed.) *Financial Incentives For Academic Quality.* New Directions for Higher Education, No.48. San Francisco: Jossey-Bass, December 1984, 29-41.

Edgerton, Russell. "Assessment is Coming: Should We Retreat or Charge?" *Higher Education and National Affairs,* (February 24, 1986) 11.

Ewell, Peter. *Information on Student Outcomes: How to Get It and How to Use It.* Boulder, CO: NCHEMS, 1983.

Ewell, Peter. *The Self-Regarding Institution: Information for Excellence.* Boulder, CO: NCHEMS, 1984.

Excellence & Equity: A Strategic Three Year Plan For Kean College of New Jersey. Kean College of New Jersey: 1985.

Fraher, Richard. "Learning a New Art: Suggestions for Beginning Teachers" in M., Gullette (ed.), *The Art and Craft of Teaching.* Harvard University Press, 1984.

Hanford, George H. Testimony prepared for delivery before the National Governors' Association Task Force on College Quality, Washington, DC, February 22, 1986, Unpublished manuscript.

Lazarus, Mitchell. "Coming to Terms With Testing," in Paul L. Houts (ed.), *The Myth of Measurability.* New York: 1977.

Northeast Missouri State University. *In Pursuit of Degrees With Integrity: A Value-Added Approach to Undergraduate Assessment.* Washington, DC: American Association of State Colleges and Universities, 1984.

Popham, James W. *Criterion-Referenced Instruction.* Belmont, CA: 1973.

Warren, Jonathan. Workshop, "Course Examination Questions in the Institutional Documentation of Student Learning," Wagner College, Staten Island, NY, February 13, 1986.

APPENDIX A

Statement of Principles

President Weiss established this task force of faculty, administrators, and students as a step in implementing the value-added assessment program detailed in the Challenge Grant Proposal. We have been charged to: 1) study approaches to value-added which are being used in other colleges, 2) identify those measurement approaches now being used at Kean, 3) disseminate information about these approaches to the Kean College community, 4) make policy recommendations for implementing a value-added program, and 5) consult with governing groups at the College to insure that appropriate decision-making channels are used to approve recommendations.

In accordance with the charge from President Weiss, the task force members will be guided in their work by the following principles:

1. The purpose of value-added assessment is to improve student learning and development by creating methods to: a) articulate the goals of each academic program, b) gain feedback on the progress towards those goals, and c) use the feedback to modify aspects of each academic program to ensure that the goals are being achieved.

2. There are many valid approaches to value-added assessment. Evaluations may be quantitative or qualitative. The faculty in each program will identify or create its own assessment instruments.

3. Value-added assessments will not be used as "gates" to control entrance to or exit from programs.

4. Value-added assessment results are solely for Kean College use; only data which are currently required and presently reported (e.g., affirmative action data, Developmental Studies results, graduate survey information, etc.) will be made available to the Department of Higher Education and other external agencies. The educational soundness of the value-added program requires that confidentiality of the data be maintained.

5. Data from value-added assessment shall not be used to make comparisons among faculty, departments, schools, or colleges.

6. Data from the value-added program shall not be used for faculty evaluation and, in particular, shall not be employed in the retention, tenure, promotion, or merit award processes.

7. The faculty will be provided with guidance, assistance, and financial support to implement each program's assessment system.

8. Value-added assessment is a long-range process. The next three years will be used to identify approaches, to study and refine them, and to implement systems for ongoing program assessment.

Principles adopted November, 1985.

APPENDIX B

TASK FORCE ON STUDENT LEARNING AND DEVELOPMENT

Robert Andrews	Asst. Professor	Communication Sciences and Educational Services
Jose Baldassini	Asst. Professor	Instruction, Curriculum and Administration
Beverly Berry	Director	E.E.O. Program
James Daly	Asst. Professor	Political Science
Catherine Dorsey-Gaines	Professor	Early Childhood and Family Studies
Doris Engelhard	Assoc. Professor	Music
Marion Epstein	Chairperson	Board of Trustees
Virginia Fitzsimons	Assoc. Professor	Nursing
Dorothy Goldberg	Professor	Mathematics
Lidija Kampa	Acting Asst. Dean	School of Natural Sciences Nursing and Mathematics
Henry Kaplowitz	Assoc. Professor	Psychology
Mary Lewis	Dean	School of Humanities, Social and Administrative Sciences
Carla Lord	Assoc. Professor	Fine Arts
Donald Lumsden	Professor	English Speech-Theatre-Media
Mary Lou Mayo	Assoc. Professor	Sociology and Social Work
Janice Murray	Assoc. Dean of Students	Division of Student Services
Daniel O'Day	Professor	English
Ruth O'Dell	Member	Board of Trustees
Peter Pezzolo	Assoc. Professor	Philosophy
Gene Rhie	Librarian 2 Asst. Professor	Library
Henry Ross	Director	Office of Institutional Research

Gary Schader	Asst. Professor	Economics/Management Science
Ana Maria Schuhmann	Acting Asst. Dean	School of Education, Technology and Related Professions
Sylvia Strauss	Assoc. Professor	History
Joseph Vitale	Director	Academic and Career Planning and Placement